Document généré le 1 nov. 2024 08:13

**Canadian Medical Education Journal**
**Revue canadienne de l'éducation médicale**

# Designing a touchless physical examination for a virtual Objective Structured Clinical Examination

# Conception d’un examen physique sans contact pour un examen clinique objectif structuré virtuel

Wassim Karkache, Samantha Halman, Christopher Tran, Rui Nie et Debra Pugh

Volume 15, numéro 2, 2024

URI : https://id.erudit.org/iderudit/1111559ar
DOI : https://doi.org/10.36834/cmej.74261

Aller au sommaire du numéro

Éditeur(s)

Canadian Medical Education Journal

ISSN

1923-1202 (numérique)

Découvrir la revue

Citer cet article

Karkache, W., Halman, S., Tran, C., Nie, R. & Pugh, D. (2024). Designing a touchless physical examination for a virtual Objective Structured Clinical Examination. *Canadian Medical Education Journal / Revue canadienne de l'éducation médicale*, *15*(2), 34–38. https://doi.org/10.36834/cmej.74261

Résumé de l'article

Objet : Compte tenu de la pandémie de COVID-19, de nombreux examens cliniques objectifs structurés (ECOS) ont été adaptés vers un format virtuel sans que l'on se questionne à savoir si les manœuvres d'examen physique peuvent ou doivent être évaluées virtuellement. Conséquemment, nous avons développé une nouvelle station d'examen physique sans contact pour un ECOS virtuel et recueilli des preuves de validité concernant son utilisation.

Méthodes : Nous avons utilisé une station d'examen physique sans contact testée dans le cadre d'un ECOS virtuel pendant lequel les résidents en médecine interne devaient verbaliser leur approche concernant l'examen physique, interpréter des images et des vidéos d'examens fournis sur demande, et poser un diagnostic. Nous avons étudié les différences de rendement en fonction de l'année de formation à l'aide de l'ANOVA. En outre, nous avons analysé les données en utilisant les éléments de la taxonomie de l'apprentissage de Bloom, c'est-à-dire la connaissance, la compréhension et la synthèse.

Résultats : Soixante-sept résidents (PGY1-3) ont participé à l'ECOS. Les scores de la station pilote étaient significativement différents entre les niveaux de formation ($F=3.936$, $p=0.024$, $\eta p2=0.11$). La corrélation totale de la station pilote (STC) était de $r=0,558$, et les corrélations question-station variaient de $r=0,115$-$0,571$, les questions les plus discriminantes étant celles qui évaluaient l'application (interprétation et synthèse) plutôt que le rappel de connaissances.

Conclusion : Cette station d'examen physique sans contact était réalisable, a présenté des caractéristiques psychométriques acceptables et a permis d'établir une discrimination entre les résidents de différents niveaux de formation.

© Wassim Karkache, Samantha Halman, Christopher Tran, Rui Nie et Debra Pugh, 2024

Ce document est protégé par la loi sur le droit d’auteur. L’utilisation des services d’Érudit (y compris la reproduction) est assujettie à sa politique d’utilisation que vous pouvez consulter en ligne.

https://apropos.erudit.org/fr/usagers/politique-dutilisation/

Cet article est diffusé et préservé par Érudit.

Érudit est un consortium interuniversitaire sans but lucratif composé de l’Université de Montréal, l’Université Laval et l’Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche.

https://www.erudit.org/fr/

Canadian Medical Education Journal

Brief Reports

# Designing a touchless physical examination for a virtual Objective Structured Clinical Examination

## Conception d'un examen physique sans contact pour un examen clinique objectif structuré virtuel

*Wassim Karkache,[1] Samantha Halman,[1,2] Christopher Tran,[1,2] Rui Nie,[3] Debra Pugh,[1,2,3]*

[1]Faculty of Medicine, The University of Ottawa, Ontario, Canada; [2]Department of Medicine, The Ottawa Hospital, Ontario, Canada; [3]Medical Council of Canada, Ontario, Canada

Correspondence to: Wassim Karkache, MD, 501 Smyth Road, Ottawa, ON K1H 8L6, phone: 613-255-7573; email: wkark064@uottawa.ca

Published ahead of issue: Mar 18, 2024; published: May 1, 2024. CMEJ 2024, 15(2) Available at https://doi.org/10.36834/cmej.74261

© 2024 Karkache, Halman, Tran, Nie, Pugh; licensee Synergies Partners. This is an Open Journal Systems article distributed under the terms of the Creative Commons Attribution License. (https://creativecommons.org/licenses/by-nc-nd/4.0) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is cited.

## Abstract

**Purpose:** Given the COVID-19 pandemic, many Objective Structured Clinical Examinations (OSCEs) have been adapted to virtual formats without addressing whether physical examination maneuvers can or should be assessed virtually. In response, we developed a novel touchless physical examination station for a virtual OSCE and gathered validity evidence for its use.

**Methods:** We used a touchless physical examination OSCE station pilot-tested in a virtual OSCE in which Internal Medicine residents had to verbalize their approach to the physical examination, interpret images and videos of findings provided upon request, and make a diagnosis. We explored differences in performance by training year using ANOVA. In addition, we analyzed data using elements of Bloom's taxonomy of learning, i.e. knowledge, understanding, and synthesis.

**Results:** Sixty-seven residents (PGY1-3) participated in the OSCE. Scores on the pilot station were significantly different between training levels ($F=3.936$, $p = 0.024$, $\eta_p^2 = 0.11$). The pilot station-total correlation (STC) was $r = 0.558$, and the item-station correlations ranged from $r = 0.115$-$0.571$, with the most discriminating items being those that assessed application of knowledge (interpretation and synthesis) rather than recall.

**Conclusion:** This touchless physical examination station was feasible, had acceptable psychometric characteristics, and discriminated between residents at different levels of training.

## Résumé

**Objet :** Compte tenu de la pandémie de COVID-19, de nombreux examens cliniques objectifs structurés (ECOS) ont été adaptés vers un format virtuel sans que l'on se questionne à savoir si les manœuvres d'examen physique peuvent ou doivent être évaluées virtuellement. Conséquemment, nous avons développé une nouvelle station d'examen physique sans contact pour un ECOS virtuel et recueilli des preuves de validité concernant son utilisation.

**Méthodes :** Nous avons utilisé une station d'examen physique sans contact testée dans le cadre d'un ECOS virtuel pendant lequel les résidents en médecine interne devaient verbaliser leur approche concernant l'examen physique, interpréter des images et des vidéos d'examens fournis sur demande, et poser un diagnostic. Nous avons étudié les différences de rendement en fonction de l'année de formation à l'aide de l'ANOVA. En outre, nous avons analysé les données en utilisant les éléments de la taxonomie de l'apprentissage de Bloom, c'est-à-dire la connaissance, la compréhension et la synthèse.

**Résultats :** Soixante-sept résidents (PGY1-3) ont participé à l'ECOS. Les scores de la station pilote étaient significativement différents entre les niveaux de formation ($F=3.936$, $p=0.024$, $\eta p2=0.11$). La corrélation totale de la station pilote (STC) était de $r=0,558$, et les corrélations question-station variaient de $r=0,115$-$0,571$, les questions les plus discriminantes étant celles qui évaluaient l'application (interprétation et synthèse) plutôt que le rappel de connaissances.

**Conclusion :** Cette station d'examen physique sans contact était réalisable, a présenté des caractéristiques psychométriques acceptables et a permis d'établir une discrimination entre les résidents de différents niveaux de formation.

# Introduction

The COVID-19 pandemic forced many institutions to transition to virtual Objective Structured Clinic Examination (OSCE) formats for the assessment of clinical skills.[1] One of the challenges with pivoting to a virtual format is how to assess physician examination (PE) skills. Thus far, approaches to assessing PE skills during virtual OSCEs have varied. One approach is to have examinees focus only on 'inspection' when examining standardized patients (SPs).[1,2] Another is to have examinees only verbalize what PE maneuvers they would want to conduct.[3] Others have removed all PE components[4] while some have incorporated multimedia to illustrate real clinical findings.[1,5] The use of high-fidelity virtual patients has also been reported.[6] Studies have reported some challenges in assessing PE skills virtually, although it is unclear if this is due to a knowledge gap, the limitations of the format, or inadequate training with the technology.[7,8]

While virtual OSCEs are feasible, there are limitations regarding what PE skills and competencies can and should be assessed with this format.[1,3] One can conceptualize a PE as four distinct competencies: selecting appropriate PE maneuvers, performing these maneuvers, identifying normal and abnormal findings, and interpreting the significance of the findings. One disadvantage of traditional OSCEs is that SPs typically do not have abnormal findings.[9] While examinees may be able to demonstrate a PE, it is not clear that they would be able to interpret findings to make a diagnosis.[9] Thus, although lower order skills such as knowledge recall are assessed (e.g. performing PE maneuvers by rote), skills related to application of knowledge, such as interpretation and synthesis (e.g. interpreting the significance of abnormal findings) may not be assessed. A virtual format may present an opportunity to target the assessment of these higher order skills, for example by incorporating multimedia depicting abnormal findings, something that can be challenging in a traditional setting.

While previous studies have looked at studying virtual physical exam stations with or without adding abnormal findings, they either did so without the learners having to identify the abnormal findings independently, there were no abnormal findings and/or there was no discrimination between the different orders of learning.[1,2,3,4,5,10,11] In this study, we developed and pilot-tested a touchless PE station within a virtual OSCE. We prompted examinees to verbalize their approach to a PE, interpret findings using images, audio, and video files, and formulate a diagnosis. We aimed to assess the feasibility of implementing a touchless PE station that assesses knowledge recall (i.e. information gathering) and application of knowledge (i.e. interpretation and synthesis) and to gather validity evidence related to scoring, generalization and extrapolation.

# Methods

The Ottawa Hospital Research Ethics Board (OHREB) reviewed this project, which received an exemption in the context of a quality initiative/program evaluation.

## Participants

At the University of Ottawa, Internal Medicine (IM) residents from post-graduate years one to three (PGY-1 to 3) participated in a mandatory formative OSCE in March 2021. As this is a progress test, residents from all years participate but the difficulty level is set at a PGY-3 level.

## OSCE format administration

We used Microsoft Teams to implement the OSCE virtually. Participants logged in remotely. The OSCE consisted of nine 12-minute stations: four structural oral and four PE stations, and one communication station. Skills assessed included knowledge, PE skills, and communication skills. There were no SPs.

We recruited physicians as examiners; they underwent training using recorded orientation sessions and detailed written instructions. For the pilot station, they received additional training on how to share PE findings (i.e. using images, audio files, videos).

We used station-specific checklists, rating scales, and examiner oral questions to score examinees.

## Pilot station development

The co-investigators developed the pilot station. The checklist consisted of 18 items, one station-specific rating scale, and a global rating scale (for standard-setting purposes). The pilot station used elements of Bloom's taxonomy of learning where we sub-categorized individual checklist items two domains: knowledge recall or application of knowledge (i.e. interpretation and synthesis).[12] Participants were unaware of item categorization.

Examinees received directions to complete a touchless PE and to describe any normal or abnormal findings. As examinees progressed through the station, the station provided them with real-time images, audio files and videos of PE findings, and expected them to interpret and

recognize their significance. In contrast, the other PE stations followed a more traditional format, i.e. examinees verbally describing PE maneuvers with examiners reporting any findings.

### Analyses

We calculated total scores for the overall OSCE using scores on all 9 stations (weighted equally). We calculated the reliability (internal consistency) of the OSCE using Cronbach's alpha. We calculated cut scores using the borderline group regression method, which incorporates a linear regression approach using the scores from all examinees.

We used univariate ANOVA to compare differences in scores between training levels for the overall OSCE and for the pilot station scores, domain sub-scores and rating scale scores. To explore pair-wise differences, we used Tukey's honestly significant difference test. We calculated partial eta squared ($\eta_p^2$) to measure effect size.[13] We hypothesized that overall and application of knowledge scores would increase with training level, while knowledge recall scores would be similar across training levels.

We calculated station-total correlations between each station and item-total correlations for each item on the pilot station. To compare total scores, domain sub-scores and rating scale score for the pilot station, we analyzed Pearson's correlation coefficients. We used SAS, version 9.4 (SAS Institute Inc) for statistical analyses.

## Results

Sixty-seven residents participated in this OSCE (25 PGY-1, 21 PGY-2, and 21 PGY-3 residents). The mean score for the overall OSCE was 66.2% ±9.2 and the cut score was 58.2. The reliability for the OSCE, measured by Cronbach's alpha, was 0.79.

The overall pass rate for the OSCE was 60% (mean=59.1 ±6.0), 90.5% (mean=67.4 ±7.7) and 100% (mean=73.6 ± 7.3) for PGY1, PGY2 and PGY3 residents, respectively. For the pilot station, the total mean score was 6.8 ±1.4 (out of 10).

Total scores on the overall OSCE were significantly different between training levels ($F=24.828$, $p < 0.001$, $\eta_p^2 = 0.437$). Post-hoc analyses revealed differences between all levels, with scores increasing significantly with each year of training. Total scores and rating scale scores on the pilot station were also significantly different between training levels ($F = 3.936$, $p = 0.024$, $\eta_p^2=0.11$; $F=11.76$, $p < 0.001$, $\eta_p^2$ 0.27, respectively) with differences between PGY-1 and PGY-2, and also between PGY-1 and PGY-3 residents.

Domain sub-score analyses revealed that residents performed better on tasks relating to application of knowledge than to knowledge recall (see Table 1). There was no significant difference in sub-scores by training level (see Table 2). When looking at individual items, however, there were differences in scores between senior (PGY-2 and 3) and junior (PGY-1) examinees for five items, all of which assessed application of knowledge. For the overall OSCE, station-total correlations ranged from 0.38 to 0.75. The pilot station station-total correlation was 0.558 and the "total but excluding pilot station" correlation was 0.443 ($p < 0.001$). This similarity in station-total correlations suggests that the pilot station was measuring similar constructs as other stations. Pilot station item-station correlations ranged from 0.12 to 0.57. The three most discriminating items assessed skills relating to application of knowledge (see Table 2).

Total score and all sub-scores significantly correlated with the global rating scale score (see Table 2).

*Table 1. Pilot Station Pass Rate and Mean Pilot Station Sub-Category Score by PGY level*

| | Pass Rate (%) | Knowledge (%) | Interpretation + Synthesis (%) |
|---|---|---|---|
| PGY1 | 40 | 57.6 | 66.5 |
| PGY2 | 76.2 | 66.2 | 73.8 |
| PGY3 | 76.2 | 66.7 | 72.0 |

*Table 2. ANOVA and Tukey's HSD Test for Total Pilot Station Score, Rating-Scale Score, and the Five Highest Item Total Correlation OSCE Items*

| Evaluated Item | F | P | Eta2 | Comparison[a] | Lower CL | Difference | Upper CL |
|---|---|---|---|---|---|---|---|
| Total | 3.94 | 0.024[g] | 0.11 | 2 - 1 | -0.06 | 0.87 | 1.79 |
| | | | | 3 - 1 | 0.05 | 0.97 | 1.89 |
| | | | | 3 - 2 | -0.86 | 0.10 | 1.06 |
| Rating-Scale | 11.76 | <0.001 | 0.27 | 2 - 1 | 0.41 | 2.25 | 4.10 |
| | | | | 3 - 1 | 1.83 | 3.68 | 5.53 |
| | | | | 3 - 2 | -0.50 | 1.43 | 3.36 |
| Q6-1[b] | 9.11 | <0.001 | 0.22 | 2 - 1 | -0.08 | 0.24 | 0.56 |
| | | | | 3 - 1 | 0.25 | 0.57 | 0.89 |
| | | | | 3 - 2 | 0 | 0.33 | 0.67 |
| Q6-2[c] | 8.06 | <0.001 | 0.20 | 2 - 1 | -0.13 | 0.18 | 0.49 |
| | | | | 3 - 1 | 0.20 | 0.51 | 0.82 |
| | | | | 3 - 2 | 0.01 | 0.33 | 0.66 |
| Q8-3[d] | 4.97 | 0.010 | 0.14 | 2 - 1 | -0.44 | -0.13 | 0.19 |
| | | | | 3 - 1 | -0.73 | -0.41 | -0.09 |
| | | | | 3 - 2 | -0.62 | -0.29 | 0.05 |
| Q9[e] | 4.08 | 0.022 | 0.11 | 2 - 1 | 0.01 | 0.35 | 0.69 |
| | | | | 3 - 1 | 0.01 | 0.35 | 0.68 |
| | | | | 3 - 2 | -0.36 | 0 | 0.36 |
| Q10[f] | 7.59 | 0.001 | 0.19 | 2 - 1 | -0.01 | 0.28 | 0.58 |
| | | | | 3 - 1 | 0.18 | 0.47 | 0.77 |
| | | | | 3 - 2 | -0.12 | 0.19 | 0.50 |

[a]*Numbers denoted year of training, i.e. 1 signifies PGY 1, et cetera*
[b]*Identify cutaneous hand findings of dermatomyositis*
[c]*Examine for torso and back cutaneous manifestations of dermatomyositis*
[d]*Recognize systolic ejection murmur in keeping with aortic stenosis*
[e]*Recognize the need to assess for lymphadenopathy*
[f]*Elaborate a differential diagnosis*
[g]*Bold font indicates statistical significance (p < 0.05)*

# Discussion

The adaptation of OSCEs to a virtual format presented an opportunity to easily incorporate technology to assess different PE skills (i.e. interpretation of abnormal findings) than are traditionally assessed. The assessment of PE skills is important because previous studies have only demonstrated a modest correlation between physical examination technique and diagnostic accuracy[9] and that trainees are often unable to detect abnormal findings in an OSCE setting after completing medical school.[14,15] Thus, traditional OSCEs, in which examinees interact with a standardized patient with no abnormal findings, may limit the conclusions that one can infer based on examinee performance. Although other studies have incorporated multi-media in their OSCEs, this is one of the few studies that we know of that has leveraged the virtual format to incorporate abnormal PE findings in an OSCE for medical residents.[16]

In addition to demonstrating feasibility, careful attention must be paid to validity evidence when adapting to a virtual OSCE format.[17] In this study, a rigorous approach to the development of the content and scoring instruments was used. Furthermore, no significant technical issues were encountered, and the overall reliability and station-total correlation were acceptable. Finally, this station was able to discriminate between learners at different stages of training.

There was an unexpected lack of difference in sub-scores by year of training which may be a consequence of only having one touchless PE station. It is also possible that there may truly be no difference in the acquisition of these skills in more senior residents. Interestingly, the most discriminating items were those that did assess knowledge application skills.

This study has several limitations: due to the nature of the OSCE being virtual, we did not assess psychomotor skills; this was a single-center study with only one touchless PE station; and we did not compare findings to those of a traditional OSCE.

# Conclusion

This study demonstrates that the incorporation of a touchless PE station in a virtual OSCE is feasible and can be used to assess knowledge application skills while discriminating between learners at different levels of training.

**Conflicts of Interest:** The authors report no conflicts of interest.
**Funding**: This study did not receive any specific grant or external funding.
**Edited by:** Cindy Schmidt (senior section editor); Marcel D'Eon (editor-in-chief)

# References

1. Hopwood J, Myers G, Sturrock A. Twelve tips for conducting a virtual OSCE. *Med Teach*. 2021;43(6):633-636. https://doi.org/10.1080/0142159X.2020.1830961
2. Blythe J, Patel NSA, Spiring W, et al. Undertaking a high stakes virtual OSCE ("VOSCE") during Covid-19. *BMC Med Educ*. 2021;21(1):221. https://doi.org/10.1186/s12909-021-02660-5
3. Craig C, Kasana N, Modi A. Virtual OSCE delivery: The way of the future?. *Med Educ*. 2020;54(12):1185-1186. https://doi.org/10.1111/medu.14286
4. Boyle JG, Colquhoun I, Noonan Z, et al. Viva la VOSCE?. *BMC Med Educ*. 2020;20(1):514. https://doi.org/10.1186/s12909-020-02444-3
5. Hytönen H, Näpänkangas R, Karaharju-Suvanto T, et al. Modification of national OSCE due to COVID-19 – Implementation and students' feedback. *Eur J Dent Educ*. 2021;25(4):679-688. https://doi.org/10.1111/eje.12646
6. Mahrous A, Alammari R, Elgreatly A. Implementing virtual OSCE using an open-source online interactive 3D library. *J Dent Educ*. 2020;85(S1):1037-1039. https://doi.org/10.1002/jdd.12449
7. Sartori DJ, Olsen S, Weinshel E, Zabar SR. Preparing trainees for telemedicine: a virtual OSCE pilot. *Med Educ*. 2019;53(5):517-518. https://doi.org/10.1111/medu.13851
8. Boardman D, Wilhite JA, Adams J, et al. Telemedicine Training in the COVID Era: Revamping a Routine OSCE to Prepare Medicine Residents for Virtual Care. *J Med Educ Curric Dev*. 2021;8:23821205211024076. https://doi.org/10.1177/23821205211024076
9. Hatala R, Issenberg B, O Kassen B, Cole G, Bacchus M, Scalese R. Assessing the Relationship between Cardiac Physical Examination Technique and Accurate Bedside Diagnosis during an Objective Structured Clinical Examination (OSCE). *Acad Med*. 2007;82(10):S26-S29. https://doi.org/10.1097/acm.0b013e31814002f1
10. Courteille O, Bergin R, Stockeld D, Ponzer S, Fors U. The use of a virtual patient case in an OSCE-based exam–a pilot study. *Med Teach*. 2008;30(3):e66-76. https://doi.org/10.1080/01421590801910216
11. Luke S, Petitt E, Tombrella J, McGoff E. Virtual evaluation of clinical competence in nurse practitioner students. *Med Sci Educ*. 2021;31(4):1267-1271. https://doi.org/10.1007/s40670-021-01312-z
12. White CB, Ross PT, Haftel HM. Assessing the assessment: are senior summative OSCEs measuring advanced knowledge, skills, and attitudes? *Acad Med.* 2008;83(12):1191-1195. https://doi.org/10.1097/acm.0b013e31818c6f6a
13. Richardson JTE. 2011. Eta squared and partial eta squared as measures of effect size in educational research. *Educ Res Rev*. 2011;6:135-147. https://doi.org/10.1016/j.edurev.2010.12.001
14. Ramani S, Ring BN, Lowe R, Hunter D. A pilot study assessing knowledge of clinical signs and physical examination skills in incoming medicine residents. *J Grad Med Educ*. 2010;2(2):232-235. https://dx.doi.org/10.4300%2FJGME-D-09-00107.1
15. Wilson BE. Performance-based assessment of internal medicine interns: evaluation of baseline clinical and communication skills. *Acad Med*. 2002;77(11):1158. https://doi.org/10.1097/00001888-200211000-00023
16. Chan SCC, Choa G, Kelly J, Maru D, Rashid MA. Implementation of virtual OSCE in health professions education: a systematic review [published online ahead of print, 2023 Apr 20]. *Med Educ*. 2023. https://doi.org/10.1111/medu.15089
17. Hess BJ, Kvern B. Using Kane's framework to build a validity argument supporting (or not) virtual OSCEs. *Med Teach*. 2021;43(9):999-1004. https://doi.org/10.1080/0142159X.2021.1910641
18. Cook DA, Brydges R, Ginsburg S, Hatala R. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Med Educ*. 2015;49(6):560-575. https://doi.org/10.1111/medu.12678